SIN'20 CONF

POLYTECH
Peter the Great
St.Petersburg Polytechnic
University

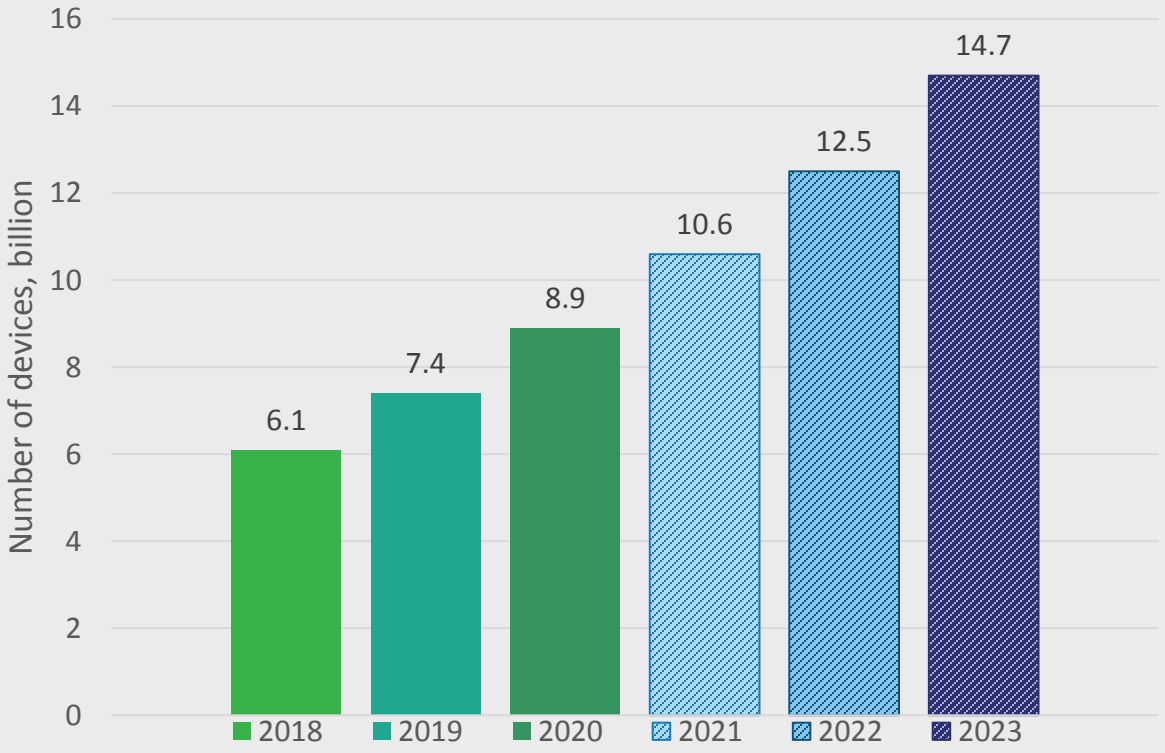
Detection of attacks on the Internet of Things based on intelligent analysis of devices functioning indicators

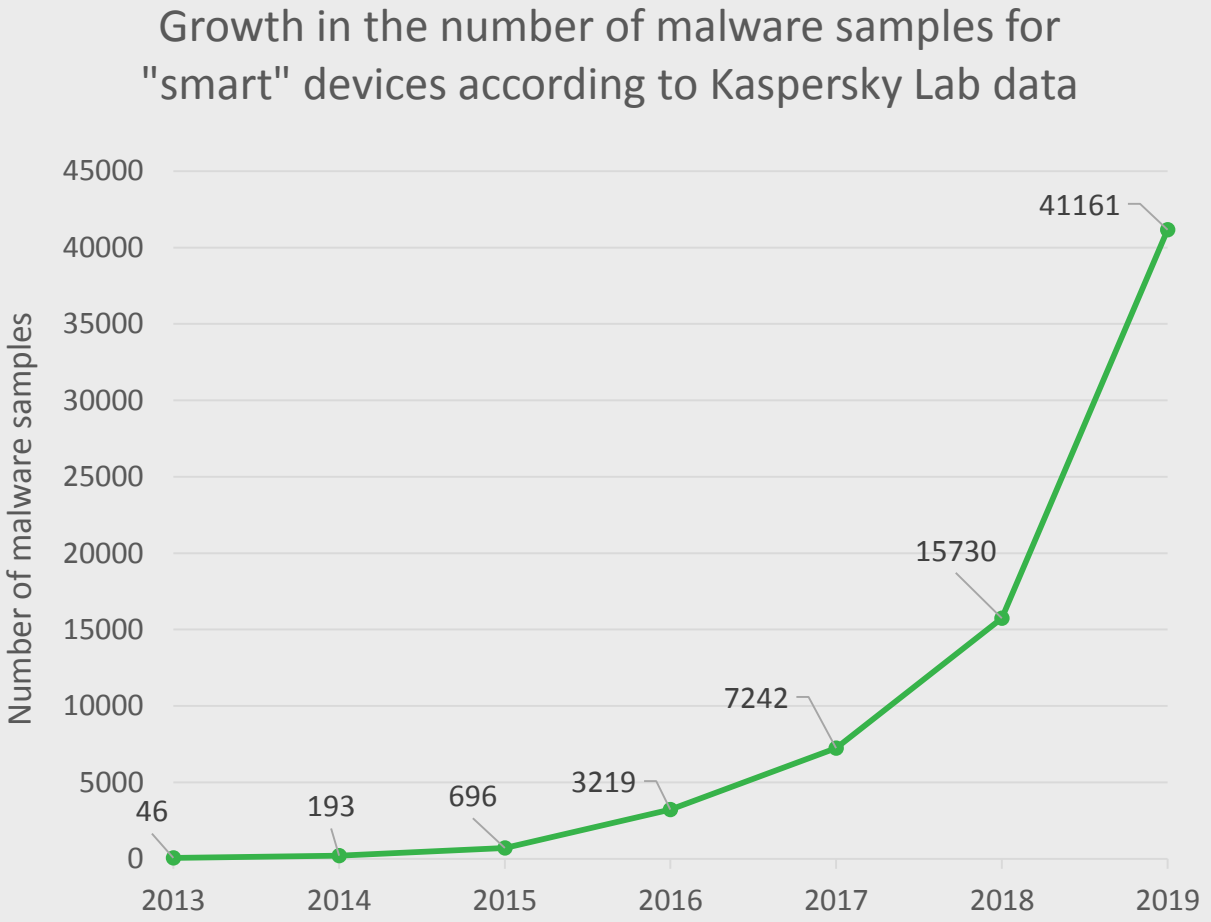
Tigran Ovasapyan
Dmitry Moskvin
Artem Tsvetkov

POLYTECH
Peter the Great St.Petersburg Polytechnic University

The relevance of research

The number of IoT devices in the world is growing every year. This is also reflected in the growth in the number of attacks.





How IoT Networks Work

The modern concept of the Internet of Things implies an interconnected, interconnected network of devices, sensors, embedded systems, mobile devices, etc. that collect, process and exchange various kinds of information.

IoT networks are used in the following areas:

- transport;
- housing and communal services;
- medicine;
- safety;
- the quality of life;
- banking;
- agriculture / animal husbandry.

Important features of IoT networks:

- distribution of network;
- processing of confidential data;
- do not interact with a person directly;
- to fully investigate attacks, it is required to emulate devices.

Comparative analysis of IoT platforms

	Contiki	Android Things	Riot	Apache Mynewt	Zephyr	TinyOS
Linux kernel based	–	+	–	–	+	–
Support for multitasking and multithreading	+/–	+	+/–	+/–	+	–
Own development tools	–	+	–	+/–	–	–
Centralized update capability	–	+/-	–	+/-	–	–
Can be used for complex tasks	–	+	–	–	+	–

POLYTECH
Peter the Great
St.Petersburg Polytechnic
University

Threats to IoT networks

Physical	Network	Software	Cryptographic
Interference with nodes	Traffic analysis	Viruses and worms	Side channel attacks
RFID radio attack	RFID spoofing	Spyware and adware	Cryptanalysis
Interfering with wireless sensor networks	RFID cloning	Trojan horses	Man in the middle attack
Injection of malicious nodes	Unauthorized RFID Access	Malicious scripts	
Physical Damage	Sinkhole attack	Denial of service attack	
Social engineering	Man in the middle attack		
Resource depletion attack	Denial of service attack		
Injection of malicious code into a site	Attack on routing mechanisms		
	Sibyl Attack		

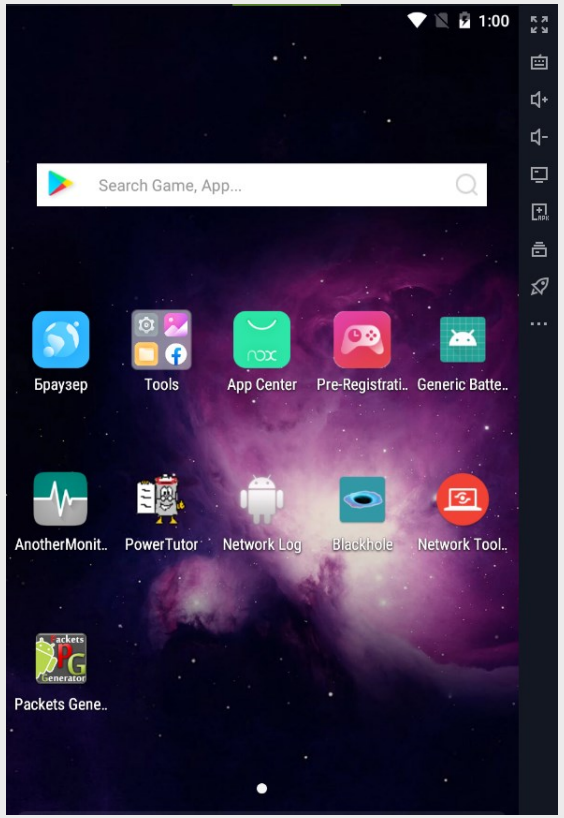
Basic methods of securing IoT devices

Static	Dynamic
– signature analysis – analysis of application components – bytecode analysis	– profiling – tracking abnormal behavior – analysis of virtual environment behavior
Disadvantage: when a new infection or attack method appears, the method may not respond to malware until it is detected and signatures for a specific vulnerability are generated	**Disadvantage:** the possibility of false positives, as well as the existing need for preliminary collection of data on the legitimate state of the system and further tuning of algorithms

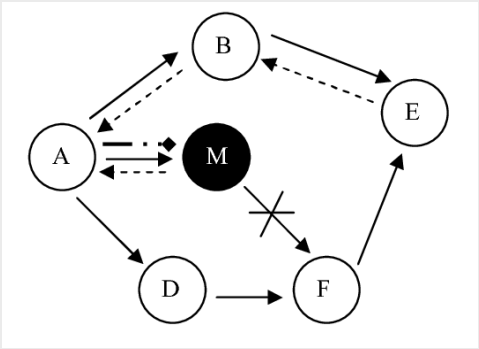
Solution for IoT devices: Leverage dynamic analysis techniques from remote IoT device data

Device emulation and attack implementation

Android Things OS is built on Android mobile OS
You can use an Android mobile device emulator to test and experiment on various attacks

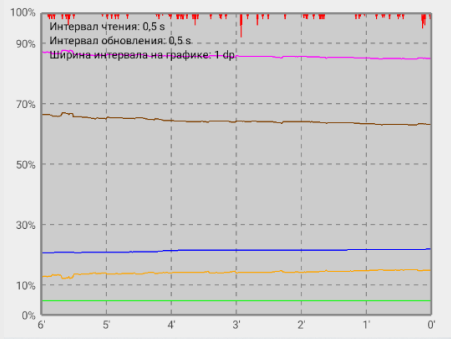


Nox Player interface



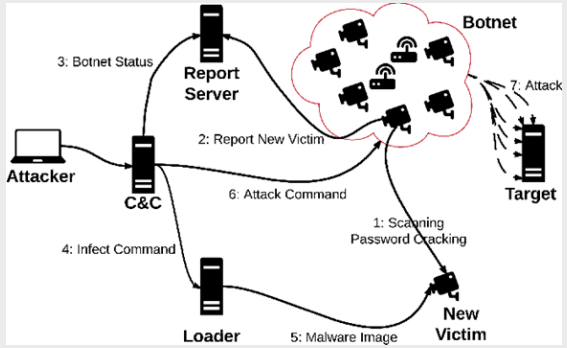
Routing attacks

Features:
- difficult to detect by standard methods
- can disable an entire network



Resource depletion attack

Features :
- difficult to detect by standard methods
- causes great harm to devices and users

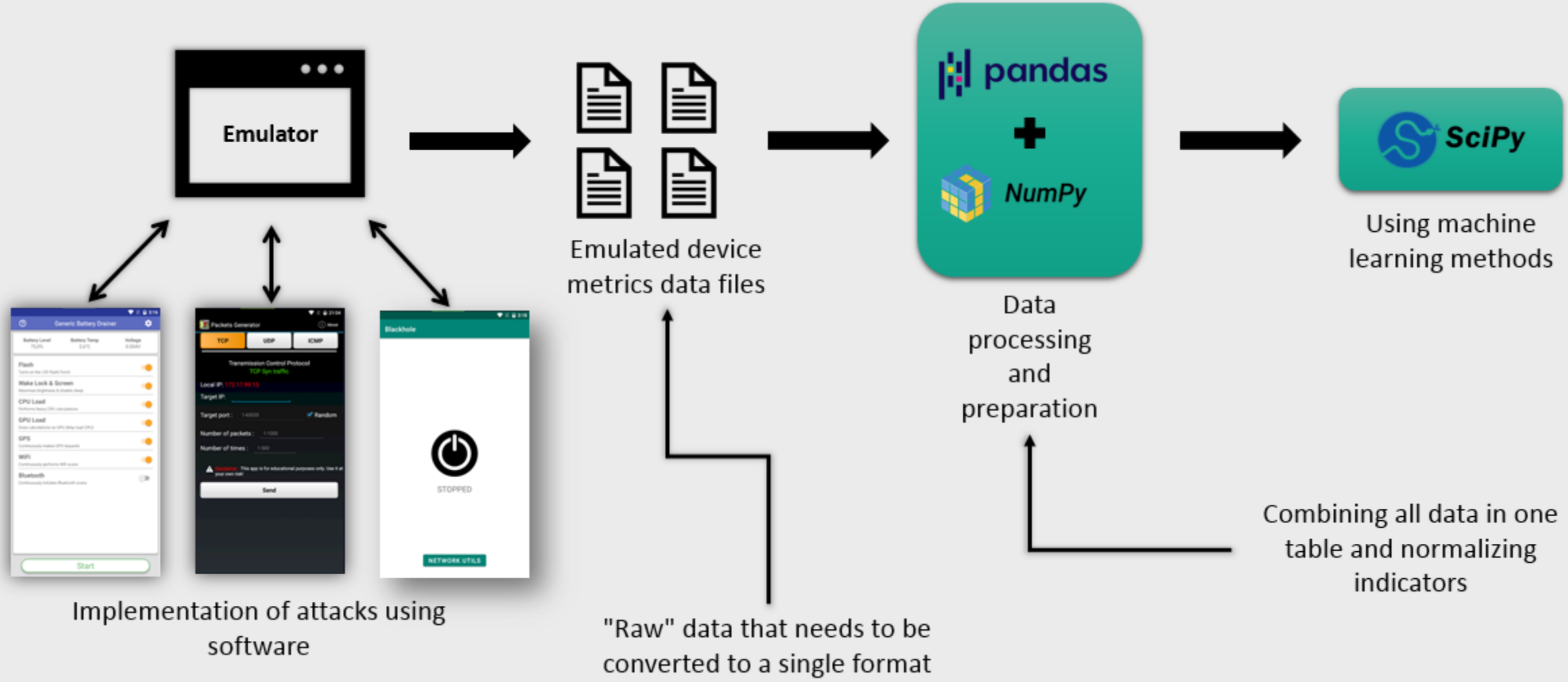


Link Noise Attack

Features :
- allows you to disrupt the entire network on which the infected host is located
- allows you to turn your device into a bot

POLYTECH
Peter the Great St.Petersburg Polytechnic University

General scheme of implementation of the method for detecting attacks in IoT networks



Collecting device metrics data to implement a method for detecting attacks in IoT networks

An example of data obtained during the emulation of attacks:

	Total CPU usage (%)	Memory used (kB)	Memory available (kB)	MemFree (kB)	Cached (kB)
0	3.000000	426680	2684152	2555620	128532
1	4.784689	424916	2685916	2557388	128528
2	2.512563	424592	2686240	2557596	128644
3	2.955665	424420	2686412	2557768	128644
4	1.515151	424344	2686488	2557844	128644
5	2.955665	424352	2686480	2557836	128644
6	2.000000	424468	2686364	2557720	128644
7	0.505050	424644	2686188	2557544	128644
8	1.015228	424468	2686364	2557720	128644
9	0.000000	424620	2686212	2557568	128644
10	1.522843	424344	2686488	2557844	128644

For each attack emulation experiment, we managed to obtain the following data from the device:

- total processor load (in percent)
- total use of RAM
- the amount of RAM available in the system
- the amount of free RAM;
- cache

Received parameters for network interfaces:

- package send interface
- packet receiving interface
- Sending IP address
- port of dispatch
- Recipient IP address
- recipient port
- package size

Training of classifiers

Three algorithms were chosen as classifiers for building models and testing:

- k-Nearest Neighbors (k-NN)
- Support Vector Machine (SVM)
- Random Forest (RF)

During the training, the following parameters were selected:

- For kNN:
  number of neighbors k ranging from 2 to 20
- For SVM:
  regularization parameter C from the set [10^-5, 10^4,… 10^5, 10^6]
- For RF:
  the number of trees n ranges from 2 to 100

Formula for the classification algorithm kNN:

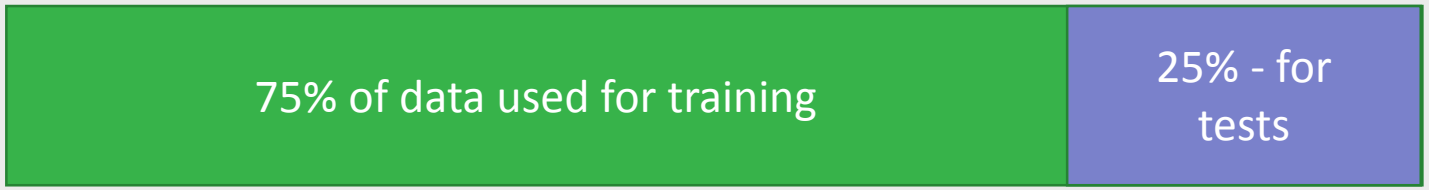
$$w(i,u) = [i \le k];$$

$$a(u; X^{\ell}, k) = \arg\max_{y \in Y} \sum_{i=1}^{k} [y_u^{(i)} = y].$$

Formula for SVM classification algorithm:

$$a(x) = sign\left(\sum_{i=1}^{\ell} \lambda_i y_i \langle x_i, x \rangle - w_0\right)$$

Sample size: 69481 records
The data is shuffled and broken up in the ratio:



Testing models and comparing learning outcomes

Results of classification by the kNN method with the number of neighbors k = 10:

Accuracy (%)	Recall (%)	Precision (%)	F-score (%)
99.8273	**99.8368**	**99.8307**	**99.8335**

Results of classification by the SVM method with the linear kernel parameter C = 10^{-2}:

Accuracy (%)	Recall (%)	Precision (%)	F-score (%)
97.2366	97.2603	97.4790	97.2233

RF classification results with number of trees n = 42:

Accuracy (%)	Recall (%)	Precision (%)	F-score (%)
97.4093	97.5220	97.6611	97.4696

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

$$precision = \frac{TP}{TP + FP} \qquad recall = \frac{TP}{TP + FN}$$

$$F = \frac{2 * precision * recall}{precision + recall}$$

Results of work

1. A study of the principles of functioning of networks of the Internet of Things and a comparative analysis of existing emulation platforms was carried out. A platform for software emulation was identified within the framework of the study.
2. Typical threats have been analyzed and actual attacks exposed to IoT networks have been identified.
3. Existing methods of detecting attacks on IoT devices are considered, their advantages and disadvantages are determined.
4. A method for detecting attacks on the Internet of Things networks using the analysis of device indicators has been developed and its effectiveness has been experimentally evaluated. The best result is achieved when using the k nearest neighbors (kNN) algorithm with the parameter k = 10.